

ML2020SPRING HW12 Report

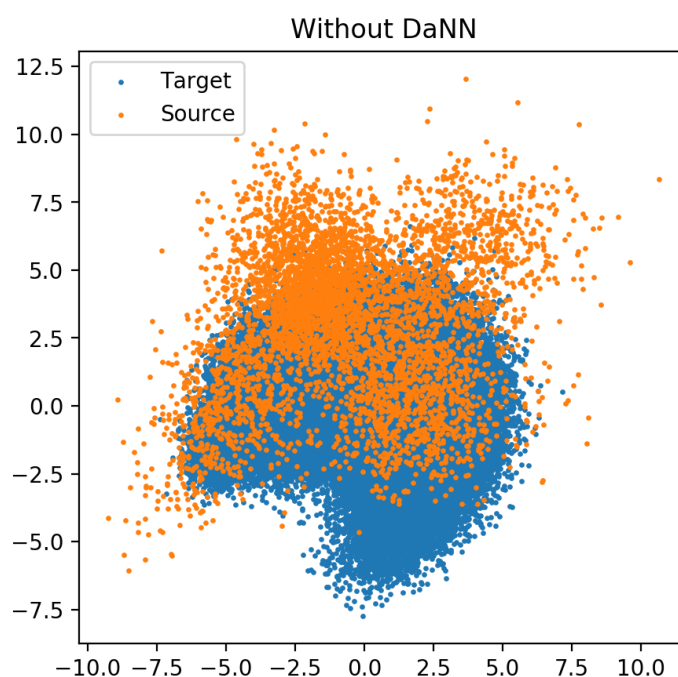
學號：R08946015 系級：資料科學碩一 姓名：陳鈞廷

1. 請描述你實作的模型架構、方法以及 accuracy 為何。其中你的方法必須為 domain adversarial training 系列(就是你的方法必須要讓輸入 training data & testing data 後的某一層輸出 distribution 要相近)。

我採用了 sample code 提供的 Domain Adversarial Training of Nerural Networks，包含了 3 個網絡模型：1. **Feature extractor** 做圖片特徵抽取，2. **Label Predictor** 做圖片分類，3. **Discriminator** 預測圖片來自哪個 domain。**Feature extractor** 我使用了 VGG11 的架構，希望多了 2 層 convolutional 可以增強特徵擷取能力，讓 **Discriminator** 可以把圖片 domain 分得更開，另外為了減緩 **Label Predictor** 的 overfitting 速度，將 fully connected layer 的 neuron 數減少。因為有些 source 的圖片包含到不重要的背景，這些背景可能會出現在做完 canny edge detection 之後的灰階圖片中，所以我在 source 的 data loader 中補上了模糊化的處理，希望可以減少這些背景資訊的影響。

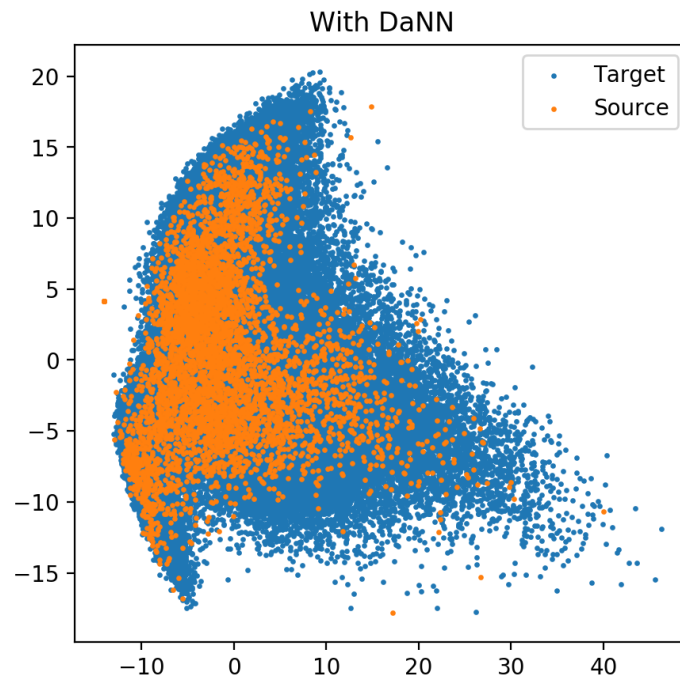
在模型的訓練上 optimizer 使用 Adam，learning rate 設定成 10^{-3} ，訓練了 500 個 epoch。因為 predict 的結果每個類別不太平衡，所以接著使用了 pseudo label semi-supervised learning：首先取出 target data 每個預測類別信心分數比較高的 2500 筆資料，把這些資料的預測類別當作標註（pseudo label），將這些資料跟 source data 一起做 feature extractor 和 label predict 的訓練，大約訓練 100 個 epoch，做完 pseudo label training 的預測結果比一開始的預測結果還要平衡一些。最後在 kaggle 的分數為 0.68。

2. 請視覺化真實圖片以及手繪圖片通過沒有使用 domain adversarial training 的 feature extractor 的 domain 分布圖。



3. 請視覺化真實圖片以及手繪圖片通過有使用 domain adversarial training 的 feature extractor 的 domain 分布圖。

2、3 兩題的圖片是將 feature extractor 輸出的 feature vector 送入 PCA 做 dimension reduction 後，得到每個 data sample 的 feature 二維投影後 plot 出來。



可以發現沒有使用 DaNN（直接使用 source data 做訓練而已）中，source data 有些點遠離了 target data 的分佈，相較之下，使用了 DaNN 的 source 分佈與 target 分佈較為接近。